

# Bayesian Hypothesis Testing for Block Sparse Signal Recovery

Mehdi Korki, *Student Member, IEEE*, Hadi Zayyani, and Jingxin Zhang

**Abstract**—This letter presents a novel **Block Bayesian Hypothesis Testing Algorithm (Block-BHTA)** for reconstructing block-sparse signals with unknown block structures. The Block-BHTA comprises the detection and recovery of the supports, and the estimation of the amplitudes of the block sparse signal. The support detection and recovery is performed using a Bayesian hypothesis testing. Then, based on the detected and reconstructed supports, the nonzero amplitudes are estimated by linear MMSE. The effectiveness of Block-BHTA is demonstrated by numerical experiments.

**Index Terms**—Block-sparse, Bayesian hypothesis testing, Bernoulli-Gaussian hidden Markov model.

## I. INTRODUCTION

Compressed sensing (CS) and sparse signal recovery aim to recover the sparse signal, a signal with only a few nonzero elements, from underdetermined systems of linear equations. In some applications, the unknown signal to be estimated has additional structure. If the structure of the signal is exploited, the better recovery performance can be achieved. A block-sparse signal, in which the nonzero samples manifest themselves as clusters, is an important structured sparsity [1]–[4]. Block-sparsity has a wide range of applications in multiband signals [5], audio signals [6], structured compressed sensing [7], and the multiple measurement vector (MMV) model [8]. The general mathematical model of the block sparse signal is

$$\mathbf{y} = \Phi \mathbf{w} + \mathbf{n}, \quad (1)$$

where  $\Phi \in \mathbb{R}^{N \times M}$  is a known measurement matrix,  $\mathbf{y} \in \mathbb{R}^N$  is the available measurement vector, and  $\mathbf{n} \in \mathbb{R}^N$  is the Gaussian corrupting noise. We aim to estimate the original unknown signal  $\mathbf{w} \in \mathbb{R}^M$ , when  $N \ll M$ , with the cluster structure

$$\mathbf{w} = [\underbrace{w_1, \dots, w_{d_1}}_{\mathbf{w}^T[1]}, \dots, \underbrace{w_{d_{g-1}+1}, \dots, w_{d_g}}_{\mathbf{w}^T[g]}]^T, \quad (2)$$

where  $\mathbf{w}[i]$  denotes the  $i$ th block with length  $d_i$  which are not necessarily identical. In the block partition (2), only  $k \ll g$  vectors  $\mathbf{w}[i]$  have nonzero Euclidean norm.

Given the *a priori* knowledge of block partition, a few algorithms such as Block-OMP [1], mixed  $\ell^2/\ell^1$  norm-minimization

M. Korki and J. Zhang are with the Department of Telecommunications, Electrical, Robotics and Biomedical Engineering, Swinburne University of Technology, Hawthorn, 3122 Australia (e-mail: mkorki@swin.edu.au; jingxinzhang@swin.edu.au).

H. Zayyani is with the Department of Electrical and Computer Engineering, Qom University of Technology, Qom, Iran (e-mail: zayyani2009@gmail.com).

[2], group LASSO [3] and model-based CoSaMP [4], work effectively in the block-sparse signal recovery. These algorithms require the knowledge of the block structure (e.g. the location and the lengths of the blocks) in (2). However, in many applications, such prior knowledge is often unavailable. Hence, devising an adaptive method for estimating the block partition and recovering the clustered-sparse signal simultaneously remains a challenge. To recover the structure-agnostic block-sparse signal, some algorithms, e.g. CluSS-MCMC [9], BMMAP-OMP [10], Block Sparse Bayesian Learning (BSBL) [11], and pattern-coupled SBL (PC-SBL) [12] have been proposed recently, which require less *a priori* information.

In this letter, we propose a novel Block Bayesian Hypothesis Testing algorithm (Block-BHTA) which uses a joint detection of the supports and estimation of the amplitudes. Block-BHTA utilizes a Bayesian hypothesis testing (BHT) for the detection and recovery of the supports. BHT was first proposed by Zayyani *et. al.* [13] in a Bayesian pursuit algorithm (BPA) for sparse representations. Recently, BHT with belief propagation has been introduced in noisy sparse recovery [14].

Inspired by BPA [13], we adopt a BHT-based approach and extend BPA to the block sparse recovery case (Block-BHTA). BPA uses the correlations between measurement vector  $\mathbf{y}$  and the columns of matrix  $\Phi$  and applies a binary BHT to obtain an activity rule in which the correlations are compared with a threshold. This activity rule is then used for the detection and recovery of the supports. Different to BPA, Block-BHTA searches for the start and termination of the blocks of the supports in the block-sparse signal  $\mathbf{w}$ . This search, performed by the BHT, leads to two ultimate activity rules where the correlations between measurement vector  $\mathbf{y}$  and the columns of matrix  $\Phi$  manifest themselves in these two activity rules. Hence, the correlations play an important role in both BPA and Block-BHTA. In these two activity rules, the correlations are compared with two simple thresholds to detect and recover the supports. Given the detected and recovered supports, Block-BHTA then uses a linear MMSE to estimate the nonzero amplitudes. Block-BHTA also uses Bernoulli-Gaussian hidden Markov model (BGHMM) [15] for the block-sparse signals. Using simple tuning updates, Block-BHTA utilizes a maximum *a posteriori* (MAP) estimation procedure to automatically learn all parameters of the statistical signal model (e.g. the variance and the elements of state-transition matrix of BGHMM). The efficiency of the proposed Block-BHTA is verified by numerical experiments.

The rest of the letter is organized as follows. In Section II, we present the signal model. In Section III, the Block-BHTA is proposed. Experimental results are presented in Section IV.

Finally, conclusions are drawn in Section V.

## II. SIGNAL MODEL

Consider the linear model of (1) as the measurement process of an underlying time- or spatial-series which is non-i.i.d and block sparse. The measurement matrix  $\Phi$  is assumed known and its columns are normalized to have unit norms. Furthermore, we model the noise in (1) as a stationary, additive white Gaussian noise (AWGN) process, with  $\mathbf{n} \sim \mathcal{N}(0, \sigma_n^2 \mathbf{I}_N)$ . To model the block-sparse sources ( $\mathbf{w}$ ), we introduce two hidden random processes,  $\mathbf{s}$  and  $\boldsymbol{\theta}$  [16], [17]. The binary vector  $\mathbf{s} \in \{0, 1\}^M$  describes the support of  $\mathbf{w}$ , denoted  $\mathcal{S}$ , while the vector  $\boldsymbol{\theta} \in \mathbb{R}^M$  represents the amplitudes of the active elements of  $\mathbf{w}$ . Hence, each element of the source vector  $\mathbf{w}$  can be characterized as

$$w_i = s_i \cdot \theta_i, \quad (3)$$

where  $s_i = 0$  gives  $w_i = 0$  for  $i \notin \mathcal{S}$  and  $s_i = 1$  gives  $w_i = \theta_i$  for  $i \in \mathcal{S}$ . In vector form, (3) can be written as  $\mathbf{w} = \mathbf{S}\boldsymbol{\theta}$ , where  $\mathbf{S} = \text{diag}(\mathbf{s}) \in \mathbb{R}^{M \times M}$ .

To model the block-sparsity of the source vector  $\mathbf{w}$ , we assume that its supports are correlated such that  $\mathbf{s}$  is a stationary first-order Markov process defined by two transition probabilities:  $p_{10} \triangleq \Pr\{s_{i+1} = 1 | s_i = 0\}$  and  $p_{01} \triangleq \Pr\{s_{i+1} = 0 | s_i = 1\}$ . Therefore, in the steady state,  $\Pr\{s_i = 0\} = p = \frac{p_{01}}{p_{10} + p_{01}}$  and  $\Pr\{s_i = 1\} = 1 - p = \frac{p_{10}}{p_{10} + p_{01}}$ , which determine the probabilities of the states in relation to the transition probabilities. The two parameters  $p$  and  $p_{10}$  completely describe the state process of the Markov chain. As a result, the remaining transition probability can be determined as  $p_{01} = \frac{p \cdot p_{10}}{(1-p)}$ . The length of the blocks of the block-sparse signal is determined by parameter  $p_{01}$ , namely, the average number of consecutive samples of ones is specified by  $1/p_{01}$  in the Markov chain.

We further assume that the amplitude vector  $\boldsymbol{\theta}$  has a Gaussian distribution with  $\boldsymbol{\theta} \sim \mathcal{N}(0, \sigma_\theta^2 \mathbf{I}_M)$ . Hence, the PDF of the  $w_i$ 's is given as

$$p(w_i) = p\delta(w_i) + (1 - p)\mathcal{N}(w_i; 0, \sigma_\theta^2), \quad (4)$$

where  $\sigma_\theta^2$  is the variance of  $\boldsymbol{\theta}$ .

Equation (4) is the well known BGHMM which is a special form of Gaussian Mixture Hidden Markov model (GHMM). The hidden variables  $s_i$  with the first-order Markov chain model in BGHMM allow implicit expression of the block-sparsity of the signal  $\mathbf{w}$  to be estimated.

## III. THE PROPOSED ALGORITHM

The proposed Block-BHTA consists of support detection and amplitude estimation. Using BHT, we first detect and recover the Block-sparse support  $\mathbf{s}$ . Then, using a linear MMSE estimator, we estimate the non-zero amplitudes of the detected supports (i.e., estimating  $\boldsymbol{\theta}$ ).

### A. Support Detection Using Bayesian Hypothesis Testing

We determine the activity of the  $j$ th element of the block-sparse signal  $\mathbf{w}$  by searching the start and termination of active blocks in  $\mathbf{w}$ . Toward that end, we assume that  $w_i$  is inactive (i.e.,  $s_i = 0$ ) and we intend to determine whether  $w_{i+1}$  is active (i.e.,  $s_{i+1} = 1$ ). This case is equivalent to searching the start of the active blocks. In the second case, we assume that  $w_i$  is active (i.e.,  $s_i = 1$ ) and we intend to determine whether  $w_{i+1}$  is inactive (i.e.,  $s_{i+1} = 0$ ). This corresponds to searching the end of active blocks. Full details are given below.

1) *Searching The Start of Active Blocks:* In order to detect the start of an active block we choose one between the hypotheses  $\mathcal{H}_{01} : s_i = 0, s_{i+1} = 1$  and  $\mathcal{H}_{00} : s_i = 0, s_{i+1} = 0$ , given the measurement vector  $\mathbf{y}$ . The Bayesian hypothesis test is

$$\hat{s}_j = \begin{cases} 1 & p(\mathcal{H}_{01,j} | \mathbf{y}) > p(\mathcal{H}_{00,j} | \mathbf{y}), \\ 0 & \text{Otherwise,} \end{cases} \quad (5)$$

where  $\mathbf{y}$  is the measurement vector. The posterior probability  $p(\mathcal{H}_{01,j} | \mathbf{y})$  is given as

$$\begin{aligned} p(\mathcal{H}_{01,j} | \mathbf{y}) &= p(s_i s_{i+1} = 01 | \mathbf{y}) = p(s_i = 0) \\ &\quad \times p(s_{i+1} = 1 | s_i = 0) \times p(\mathbf{y} | s_i s_{i+1} = 01) \\ &= p \times p_{10} \times p(\mathbf{y} | s_i s_{i+1} = 01), \end{aligned} \quad (6)$$

where  $\mathbf{y} = \sum_{j=1, j \neq i}^M \varphi_j w_j + \mathbf{n}$  and  $\varphi_j$  represents the  $j$ th column of matrix  $\Phi$ . Similarly, the posterior probability  $p(\mathcal{H}_{00,j} | \mathbf{y})$  is given by

$$\begin{aligned} p(\mathcal{H}_{00,j} | \mathbf{y}) &= p(s_i s_{i+1} = 00 | \mathbf{y}) = p(s_i = 0) \\ &\quad \times p(s_{i+1} = 0 | s_i = 0) \times p(\mathbf{y} | s_i s_{i+1} = 00) \\ &= p \times p_{00} \times p(\mathbf{y} | s_i s_{i+1} = 00), \end{aligned} \quad (7)$$

where  $p_{00} = p(s_{i+1} = 0 | s_i = 0) = 1 - p_{10}$  and  $\mathbf{y} = \sum_{j=1, j \neq i, i+1}^M \varphi_j w_j + \mathbf{n}$ . Hence, from (5)-(7), the activity rule for  $w_{i+1}$  is

$$p_{10} \times p(\mathbf{y} | s_i s_{i+1} = 01) > p_{00} \times p(\mathbf{y} | s_i s_{i+1} = 00). \quad (8)$$

Assume that we have all the estimates of  $w_j$  except for  $j \neq i+1$  and we intend to estimate  $w_{i+1}$ . We have

$$p(\mathbf{y} | s_i s_{i+1} = 00) = \frac{\exp\left(-\frac{1}{2\sigma_n^2} \left\| \mathbf{y} - \sum_{j=1, j \neq i, i+1}^M \varphi_j w_j \right\|_2^2\right)}{\sqrt{(2\pi\sigma_n^2)^N}}. \quad (9)$$

When  $s_i s_{i+1} = 01$ , we have  $\mathbf{y} = \sum_{j=1, j \neq i, i+1}^M \varphi_j w_j + \varphi_{i+1} w_{i+1} + \mathbf{n} = \sum_{j=1, j \neq i, i+1}^M \varphi_j w_j + \mathbf{n}'$ , where  $\mathbf{n}' = \varphi_{i+1} w_{i+1} + \mathbf{n}$ . Hence, the likelihood  $p(\mathbf{y} | s_i s_{i+1} = 01)$  is a multivariate Gaussian with its mean and covariance given respectively by

$$\boldsymbol{\mu}_{\mathbf{y}} = \sum_{j=1, j \neq i, i+1}^M \varphi_j w_j, \quad (10)$$

$$\boldsymbol{\Sigma}_{\mathbf{y}} = \text{Cov}(\mathbf{n}') = \sigma_n^2 \mathbf{I}_N + \sigma_\theta^2 \varphi_{i+1} \varphi_{i+1}^T. \quad (11)$$

Therefore, we can write the likelihood function as

$$p(\mathbf{y} | s_i s_{i+1} = 01) = \frac{\exp\left(-\frac{1}{2}(\mathbf{y} - \boldsymbol{\mu}_y)^T \boldsymbol{\Sigma}_y^{-1} (\mathbf{y} - \boldsymbol{\mu}_y)\right)}{\sqrt{(2\pi)^N \det(\boldsymbol{\Sigma}_y)}}. \quad (12)$$

Using the matrix inversion lemma ([18], p. 571), we can express  $\boldsymbol{\Sigma}_y^{-1}$  as

$$\boldsymbol{\Sigma}_y^{-1} = \sigma_n^{-2} \mathbf{I}_N - \frac{\boldsymbol{\varphi}_{i+1} \boldsymbol{\varphi}_{i+1}^T \sigma_n^{-2}}{1 + \left(\frac{\sigma_n}{\sigma_\theta}\right)^2}. \quad (13)$$

The determinant of  $\boldsymbol{\Sigma}_y$  can be calculated as<sup>1</sup>

$$\begin{aligned} \det(\boldsymbol{\Sigma}_y) &= (\sigma_\theta)^{2N} \det\left(\frac{\sigma_n^2}{\sigma_\theta^2} \mathbf{I}_N + \boldsymbol{\varphi}_{i+1} \boldsymbol{\varphi}_{i+1}^T\right) \\ &= (\sigma_\theta)^{2N} \left(1 + \boldsymbol{\varphi}_{i+1}^T \frac{\sigma_n^2}{\sigma_\theta^2} \boldsymbol{\varphi}_{i+1}\right) \det\left(\frac{\sigma_n^2}{\sigma_\theta^2} \mathbf{I}_N\right) \\ &= (\sigma_\theta)^{2N} \left(1 + \frac{\sigma_\theta^2}{\sigma_n^2} \boldsymbol{\varphi}_{i+1}^T \boldsymbol{\varphi}_{i+1}\right). \end{aligned} \quad (14)$$

Using (9)-(14), the Bayesian hypothesis test in (8) can be simplified to give the final activity rule for  $w_{i+1}$  as

$$\text{Activity}_{\text{START}}(w_{i+1}) \triangleq \mathbf{x}^T \boldsymbol{\varphi}_{i+1} \boldsymbol{\varphi}_{i+1}^T \mathbf{x} > \text{Th}_{1,i+1}, \quad (15)$$

where  $\text{Th}_{1,i+1}$  is defined as

$$\text{Th}_{1,i+1} \triangleq 2\sigma_n^2 \left(1 + \frac{\sigma_n^2}{\sigma_\theta^2}\right) \ln\left(\frac{p_{00}}{p_{10}} \sqrt{\left(1 + \frac{\sigma_n^2}{\sigma_\theta^2}\right)}\right), \quad (16)$$

and  $\mathbf{x} = \mathbf{y} - \boldsymbol{\Phi} \mathbf{w} - \boldsymbol{\varphi}_i w_i - \boldsymbol{\varphi}_{i+1} w_{i+1}$ . It is seen that in the activity rule  $\text{Activity}_{\text{START}}(w_{i+1})$  in (15) the correlation between the columns of matrix  $\boldsymbol{\Phi}$  and measurement vector  $\mathbf{x}$  decides between  $\mathcal{H}_{01}$  and  $\mathcal{H}_{00}$ .

2) *Searching The Termination of Active Blocks:* The detection of the end of an active block is performed by choosing one between the hypotheses  $\mathcal{H}_{10} : s_i = 1, s_{i+1} = 0$  and  $\mathcal{H}_{11} : s_i = 1, s_{i+1} = 1$ , given the measurement vector  $\mathbf{y}$ . The Bayesian hypothesis test is given as

$$\hat{s}_j = \begin{cases} 0 & p(\mathcal{H}_{10,j} | \mathbf{y}) > p(\mathcal{H}_{11,j} | \mathbf{y}), \\ 1 & \text{Otherwise.} \end{cases} \quad (17)$$

Similar to (6), we have  $p(\mathcal{H}_{10,j} | \mathbf{y}) = (1 - p) \times p_{01} \times p(\mathbf{y} | s_i s_{i+1} = 10)$ , where  $\mathbf{y} = \sum_{j=1, j \neq i+1}^M \boldsymbol{\varphi}_j w_j + \mathbf{n}$ . Likewise,  $p(\mathcal{H}_{11,j} | \mathbf{y}) = (1 - p) \times p_{11} \times p(\mathbf{y} | s_i s_{i+1} = 11)$ , where  $p_{11} = p(s_{i+1} = 1 | s_i = 1) = 1 - p_{01}$  and  $\mathbf{y} = \sum_{j=1}^M \boldsymbol{\varphi}_j w_j + \mathbf{n}$ . Therefore, we have the following inactivity rule for  $w_{i+1}$

$$p_{01} \times p(\mathbf{y} | s_i s_{i+1} = 10) > p_{11} \times p(\mathbf{y} | s_i s_{i+1} = 11). \quad (18)$$

Similar to (9), the likelihood function  $p(\mathbf{y} | s_i s_{i+1} = 10)$  is calculated as

$$p(\mathbf{y} | s_i s_{i+1} = 10) = \frac{\exp\left(-\frac{1}{2\sigma_n^2} \left\| \mathbf{y} - \sum_{j=1, j \neq i+1}^M \boldsymbol{\varphi}_j w_j \right\|_2^2\right)}{\sqrt{(2\pi\sigma_n^2)^N}}. \quad (19)$$

<sup>1</sup>We have used matrix determinant lemma, i.e.  $\det(\mathbf{A} + \mathbf{u}\mathbf{v}^T) = (1 + \mathbf{v}^T \mathbf{A}^{-1} \mathbf{u}) \det(\mathbf{A})$ , where  $\mathbf{A}$  is an invertible square matrix and  $\mathbf{u}, \mathbf{v}$  are column vectors.

Also, given  $s_i s_{i+1} = 11$ ,  $\mathbf{y} = \sum_{j=1, j \neq i+1}^M \boldsymbol{\varphi}_j w_j + \boldsymbol{\varphi}_{i+1} w_{i+1} + \mathbf{n} = \sum_{j=1, j \neq i+1}^M \boldsymbol{\varphi}_j w_j + \mathbf{n}'$ , where  $\mathbf{n}' = \boldsymbol{\varphi}_{i+1} w_{i+1} + \mathbf{n}$ . Hence, the likelihood  $p(\mathbf{y} | s_i s_{i+1} = 11)$  is a multivariate Gaussian with its covariance given by (11) and its mean by

$$\boldsymbol{\mu}'_y = \sum_{j=1, j \neq i+1}^M \boldsymbol{\varphi}_j w_j. \quad (20)$$

Also, the likelihood function  $p(\mathbf{y} | s_i s_{i+1} = 11)$  can be evaluated as

$$p(\mathbf{y} | s_i s_{i+1} = 11) = \frac{\exp\left(-\frac{1}{2}(\mathbf{y} - \boldsymbol{\mu}'_y)^T \boldsymbol{\Sigma}_y^{-1} (\mathbf{y} - \boldsymbol{\mu}'_y)\right)}{\sqrt{(2\pi)^N \det(\boldsymbol{\Sigma}_y)}}, \quad (21)$$

where  $\boldsymbol{\Sigma}_y^{-1}$  and  $\det(\boldsymbol{\Sigma}_y)$  are given in (13) and (14), respectively. Substituting (19) and (21) in (18) and using (20), the final inactivity rule for  $w_{i+1}$  can be expressed as

$$\text{Inactivity}_{\text{END}}(w_{i+1}) \triangleq \mathbf{z}^T \boldsymbol{\varphi}_{i+1} \boldsymbol{\varphi}_{i+1}^T \mathbf{z} > \text{Th}_{2,i+1}, \quad (22)$$

where  $\text{Th}_{2,i+1}$  is defined as

$$\text{Th}_{2,i+1} \triangleq 2\sigma_n^2 \left(1 + \frac{\sigma_n^2}{\sigma_\theta^2}\right) \ln\left(\frac{p_{01}}{p_{11}} \sqrt{\left(1 + \frac{\sigma_n^2}{\sigma_\theta^2}\right)}\right), \quad (23)$$

and  $\mathbf{z} = \mathbf{y} - \boldsymbol{\Phi} \mathbf{w} - \boldsymbol{\varphi}_{i+1} w_{i+1}$ .

Also, the estimates of the unknown parameters  $\sigma_n$ ,  $\sigma_\theta$ ,  $p$ ,  $p_{10}$ , and  $p_{01}$  in (16) and (23) are given by the following simple updates [19]

$$\hat{\sigma}_n = \frac{\|\mathbf{y} - \boldsymbol{\Phi} \hat{\mathbf{w}}\|_2}{\sqrt{N}}, \hat{\sigma}_\theta = \sqrt{\frac{N \mathbb{E}(y_j^2)}{M(1 - \hat{p})}}, \hat{p} = \frac{\|\mathbf{s}\|_0}{M}, \quad (24)$$

$$\hat{p}_{10} = \frac{\sum_{i=1}^{M-1} s_{i+1} (1 - s_i)}{\sum_{i=1}^{M-1} (1 - s_i)}, \hat{p}_{01} = \frac{\sum_{i=1}^{M-1} s_i (1 - s_{i+1})}{\sum_{i=1}^{M-1} s_i}, \quad (25)$$

where  $\mathbb{E}(\cdot)$  represents the expectation of a random variable.

## B. Amplitude Estimation Using Linear MMSE

Given the detection and recovery information of the binary support vector  $\mathbf{s}$  by BHT, we complete the estimation of the original unknown signal  $\mathbf{w}$  by estimating the amplitude samples of the  $\boldsymbol{\theta}$  vector.

Based on the detected vector  $\mathbf{s}$ , denoted by  $\hat{\mathbf{s}}$ , we obtain the linear MMSE estimate ([18], p. 364) of  $\boldsymbol{\theta}$  (denoted by  $\hat{\boldsymbol{\theta}}$ ) which is given as

$$\hat{\boldsymbol{\theta}} = \sigma_\theta^2 \hat{\mathbf{S}} \boldsymbol{\Phi}^T \left(\sigma_n^2 \mathbf{I}_N + \sigma_\theta^2 \boldsymbol{\Phi} \hat{\mathbf{S}} \boldsymbol{\Phi}^T\right)^{-1} \mathbf{y} \quad (26)$$

where  $\hat{\mathbf{S}} = \text{diag}(\hat{\mathbf{s}})$ .

**Algorithm 1** provides a pseudo-code implementation of our proposed Block-BHTA that gives all steps in the algorithm including BHT support detection and amplitude estimation.

**Algorithm 1** The overall Block-BHTA estimation.

---

**Input:**  $\mathbf{y}$ ,  $\Phi$ ,  $k_{max}$ , and  $\epsilon$  Initialize: Choose  $p^{(0)} \in [0.5, 1]$ ,  $\sigma_\theta^{(0)} = \sqrt{\frac{N\mathbb{E}(y^2)}{M(1-p)}}$ ,  $\sigma_n^{(0)} = \sigma_\theta^{(0)}/5$ ,  $\mathbf{w}^{(0)} = \Phi^T(\Phi\Phi^T)^{-1}\mathbf{y}$ . set difference = 1,  $k = 0$ .

1: **while** (difference >  $\epsilon$  and  $k < k_{max}$ ) **do**

2:   **BHT-detection:**

3:    **for**  $i = 0, \dots, M - 1$  **do**

4:     **if**  $\text{Activity}_{\text{START}}(w_{i+1}) > \text{Th}_{1,i+1}$  in (15) **then**

5:       set  $s_i = 1$ ,

6:     **else if**  $\text{Inactivity}_{\text{END}}(w_{i+1}) > \text{Th}_{2,i+1}$  in (22) **then**

7:       set  $s_i = 0$ ,

8:     **end if**

9:   **end for**

10:   **LMMSE estimation:**  $\mathbf{S}^{(k)} = \text{diag}(\mathbf{s}^{(k)})$ ,

11:    $\theta^{(k)} = \sigma_\theta^2 \hat{\mathbf{S}} \Phi^T (\sigma_n^2 \mathbf{I}_N + \sigma_\theta^2 \Phi \hat{\mathbf{S}} \Phi^T)^{-1} \mathbf{y}$ ,

12:   **Parameter Estimation:** using (24) and (25)

13:    $\mathbf{w}^{(k)} = \mathbf{s}^{(k)} \odot \theta^{(k)}$ .

14:   Compute the difference  $\triangleq \frac{\|\mathbf{w}^{(k+1)} - \mathbf{w}^{(k)}\|_2}{\|\mathbf{w}^{(k+1)}\|_2}$ ,  $k \leftarrow k + 1$

15: **end while**

**Output:**  $\hat{\mathbf{w}} = \mathbf{w}^{(k)}$

---

**IV. SIMULATION RESULTS**

This section presents the experimental results to demonstrate the performance of the Block-BHTA. Two experimental results are presented in this section. First, we compare the performance of the proposed Block-BHTA with that of BPA [13] versus SNR. Second, we evaluate the performance of Block-BHTA versus number of nonzero blocks and compare the performance with some block-sparse signal reconstruction algorithms.

All the experiments are conducted for 400 independent simulation runs. In each simulation run, the elements of the matrix  $\Phi$  are chosen from a uniform distribution in [-1,1] with columns normalized to unit  $\ell^2$ -norm. The Block-sparse sources  $\mathbf{w}_{\text{gen}}$  are synthetically generated using BGHMM in (4) which is based on Markov chain process. Unless otherwise stated, in all experiments  $p = 0.9$ ,  $p_{01} = 0.09$  and  $\sigma_\theta = 1$  which are the parameters of BGHMM. The measurement vector  $\mathbf{y}$  is constructed by  $\mathbf{y} = \Phi \mathbf{w}_{\text{gen}} + \mathbf{n}$ , where  $\mathbf{n}$  is zero-mean AWGN with a variance tuned to a specified value of SNR which is defined as

$$\text{SNR (dB)} \triangleq 20 \log_{10} (\|\Phi \mathbf{w}_{\text{gen}}\|_2 / \|\mathbf{n}\|_2). \quad (27)$$

We use the Normalized Mean Square Error (NMSE (dB)) as a performance metric, defined by  $\text{NMSE}(\text{dB}) \triangleq 10 \log_{10} (\|\hat{\mathbf{w}} - \mathbf{w}_{\text{gen}}\|_2^2 / \|\mathbf{w}_{\text{gen}}\|_2^2)$ , where  $\hat{\mathbf{w}}$  is the estimate of the true signal  $\mathbf{w}_{\text{gen}}$ . We compare the Block-BHTA and BPA at different noise levels. In this experiment  $N = 192$  and  $M = 512$ . We add the Gaussian white noise so that SNR, defined in (27), varies between 10 dB and 30 dB for each generated signal.

Figure 1 shows the NMSE (dB) versus SNR for both Block-BHTA and BPA. It is seen that Block-BHTA exhibit significant performance gain (almost 5 dB) over BPA.

In the second experiment, we examine the influence of the block size and the number of blocks on the estimation performance of the Block-BHTA where the block partition is unknown. Towards that end, we set up a simulation to compare the Block-BHTA with some recently developed algorithms for

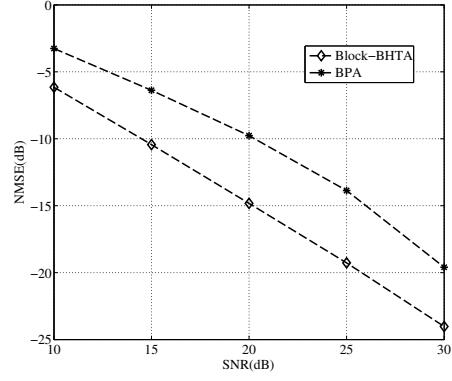


Fig. 1. NMSE (dB) versus SNR for Block-BHTA and BPA. The results are averaged over 400 trials.

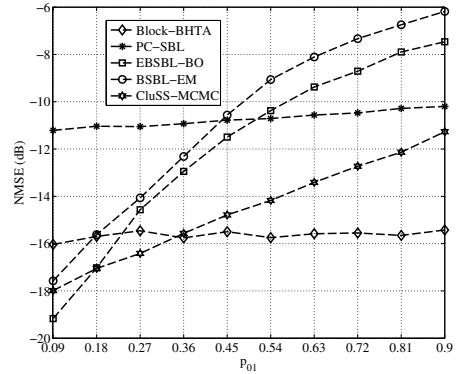


Fig. 2. NMSE versus  $p_{01}$  for Block-BHTA and other algorithms. The results are averaged over 400 trials.

block sparse signal reconstruction, such as the block sparse Bayesian learning algorithm (BSBL) [11], the expanded block sparse Bayesian learning algorithm (EBSBL) [11], the cluster-structured MCMC algorithm (CluSS-MCMC) [9], and the pattern-coupled sparse Bayesian learning algorithm (PC-SBL) [12]. The size of matrix  $\Phi$  is  $256 \times 512$ , SNR = 15dB, and  $\sigma_\theta = 1$ . Recall from Section II that the block size and the number of blocks of  $\mathbf{w}$  are proportional to  $1/p_{01}$ . That is, when  $p_{01}$  is small  $\mathbf{w}$  comprises small number of blocks with big sizes and vice versa. Hence, we vary the value of  $p_{01}$  between 0.09 and 0.9 to obtain the NMSE (dB) for various algorithms. The results of NMSE (dB) versus  $p_{01}$  is shown in Fig. 2. As seen from the figure, for  $p_{01} \geq 0.36$  the Block-BHTA outperforms all other algorithms.

**V. CONCLUSION**

This letter has presented a novel Block-BHTA to recover the block-sparse signals whose structure of block sparsity is completely unknown. The proposed Block-BHTA uses a Bayesian hypothesis testing to detect and recover the support of the block sparse signal. For amplitude recovery, Block-BHTA utilizes a linear MMSE to estimate the nonzero amplitudes of the detected supports. Simulation results demonstrate that Block-BHTA outperforms the BPA by almost 5 dB performance gain. The Block-BHTA also outperforms many state-of-the-art algorithms when the block-sparse signal comprises a large number of blocks with short lengths.

## REFERENCES

- [1] Y. C. Eldar, P. Kuppinger, and H. Boleskei, "Block-sparse signals: Uncertainty relations and efficient recovery," *Signal Processing, IEEE Transactions on*, vol. 58, no. 6, pp. 3042–3054, 2010.
- [2] Y. C. Eldar and M. Mishali, "Robust recovery of signals from a structured union of subspaces," *Information Theory, IEEE Transactions on*, vol. 55, no. 11, pp. 5302–5316, 2009.
- [3] M. Yuan and Y. Lin, "Model selection and estimation in regression with grouped variables," *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, vol. 68, no. 1, pp. 49–67, 2006.
- [4] R. G. Baraniuk, V. Cevher, M. F. Duarte, and C. Hegde, "Model-based compressive sensing," *Information Theory, IEEE Transactions on*, vol. 56, no. 4, pp. 1982–2001, 2010.
- [5] M. Mishali and Y. C. Eldar, "Blind multiband signal reconstruction: Compressed sensing for analog signals," *Signal Processing, IEEE Transactions on*, vol. 57, no. 3, pp. 993–1009, 2009.
- [6] R. Gribonval and E. Bacry, "Harmonic decomposition of audio signals with matching pursuit," *Signal Processing, IEEE Transactions on*, vol. 51, no. 1, pp. 101–111, 2003.
- [7] M. F. Duarte and Y. C. Eldar, "Structured compressed sensing: From theory to applications," *Signal Processing, IEEE Transactions on*, vol. 59, no. 9, pp. 4053–4085, 2011.
- [8] Z. Zhilin and B. D. Rao, "Sparse signal recovery with temporally correlated source vectors using sparse bayesian learning," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 5, no. 5, pp. 912–926, 2011.
- [9] L. Yu, H. Sun, J. P. Barbot, and G. Zheng, "Bayesian compressive sensing for cluster structured sparse signals," *Signal Processing*, vol. 92, no. 1, pp. 259–269, 2012.
- [10] T. Peleg, Y. C. Eldar, and M. Elad, "Exploiting statistical dependencies in sparse representations for signal recovery," *Signal Processing, IEEE Transactions on*, vol. 60, no. 5, pp. 2286–2303, 2012.
- [11] Z. Zhang and B. D. Rao, "Extension of SBL algorithms for the recovery of block sparse signals with intra-block correlation," *Signal Processing, IEEE Transactions on*, vol. 61, no. 8, pp. 2009–2015, 2013.
- [12] J. Fang, Y. Shen, H. Li, and P. Wang, "Pattern-coupled sparse bayesian learning for recovery of block-sparse signals," *Signal Processing, IEEE Transactions on*, vol. 63, no. 2, pp. 360–372, 2015.
- [13] H. Zayyani, M. Babaie-Zadeh, and C. Jutten, "Bayesian pursuit algorithm for sparse representation," in *Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on*, April 2009, pp. 1549–1552.
- [14] J. Kang, H.-N. Lee, and K. Kim, "Bayesian hypothesis test using nonparametric belief propagation for noisy sparse recovery," *Signal Processing, IEEE Transactions on*, vol. 63, no. 4, pp. 935–948, 2015.
- [15] M. Zimmermann and K. Dostert, "Analysis and modeling of impulsive noise in broad-band powerline communications," *Electromagnetic Compatibility, IEEE Transactions on*, vol. 44, no. 1, pp. 249–258, 2002.
- [16] J. Ziniel and P. Schniter, "Dynamic compressive sensing of time-varying signals via approximate message passing," *Signal Processing, IEEE Transactions on*, vol. 61, no. 21, pp. 5270–5284, 2013.
- [17] H. Zayyani, M. Babaie-Zadeh, and C. Jutten, "An iterative bayesian algorithm for sparse component analysis in presence of noise," *Signal Processing, IEEE Transactions on*, vol. 57, no. 11, pp. 4378–4390, 2009.
- [18] S. M. Kay, *Fundamentals of Statistical Signal Processing Volume II: Estimation Theory*. New Jersey, U.S.A: Englewood Cliffs, 1993.
- [19] M. Korki, J. Zhang, C. Zhang, and H. Zayyani, "An iterative bayesian algorithm for block-sparse signal reconstruction," in *Acoustics, Speech and Signal Processing (ICASSP), 2015 IEEE International Conference on*.